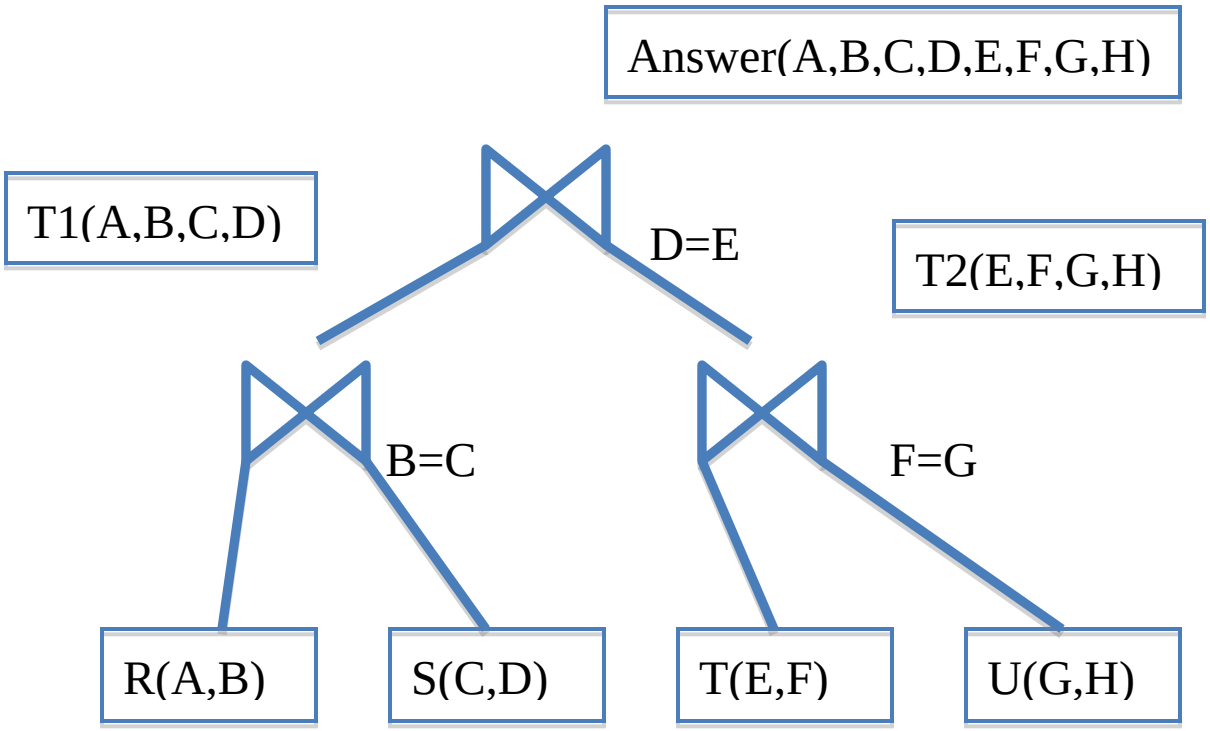
Section 9

1. Consider the query R(A,B) join S(C,D) join T(E,F) (the join condition is B=C and D=E). Suppose M =100, and B(R) = 30, B(S) = 200, B(T) = 60, B(R join S) =80, B(S join T) = 50. Design an optimal query plan that uses only main-memory hash join algorithms. Your plan may store intermediate results to disk if necessary.

2. Consider the algebra plan below. Each of the joint operators is a main memory hash join algorithm, using the Open( ), GetNext( ), Close( ) interface. Assuming that all joins are pipelining, show the execution steps for computing the entire join.



Where R, S, T, U have the following content:

R

A1	B
A2	B

S

B	D1
B	D2

T

D1	F
D2	F

U

F	H1
F	H2

(b) **[10 points]** Consider the following query, where $\bowtie$ denotes the natural join:

$$R(A,B) \bowtie S(B,C) \bowtie T(C,D) \bowtie U(D,E)$$

Here we only consider left linear plans

i. How many different left linear plans exist for this query ?

ii. Show two different left linear plans without cartesian products.

iii. How many different plans without cartesian product exists for this query ?